4.5 Design Solutions for the Bulldozer 32nm SOI 2-Core Processor Module in an 8-Core CPU

Tim Fischer[1], Srikanth Arekapudi[2], Eric Busta[1], Carl Dietz[3], Michael Golden[2], Scott Hilker[2], Aaron Horiuchi[1], Kevin A. Hurd[1], Dave Johnson[1], Hugh McIntyre[2], Samuel Naffziger[1], James Vinh[2], Jonathan White[4], Kathryn Wilcox[4]

[1]AMD, Fort Collins, CO,
[2]AMD, Sunnyvale, CA,
[3]AMD, Austin, TX,
[4]AMD, Boxborough, MA

AMD's 2-core "Bulldozer" module contains 213 million transistors in an 11-metal layer 32nm HKMG SOI CMOS process and is designed to operate from 0.8 to 1.3V. This new micro-architecture [1] improves performance and frequency while reducing area and power compared to a previous AMD x86-64 CPU in the same process [2]. To achieve these goals, the design reduced the number of FO4 inverter delays/cycle by more than 20%, achieving higher frequencies in the same power envelope even with increased core counts. The 2-core CPU module area (including 2MB L2 cache) is 30.9mm^2 (Fig. 4.5.7).

The module design contains 84 unique custom macros and 317,000 scannable flops. Module-level VSS power gating (CC6) reduces leakage power by 95% when both cores are idle [2]. Transistor Vts across the design are mostly regular (47%) and long-channel regular (46%).

The Bulldozer micro-architecture is cycle-based, using soft-edge flip-flops (SEF) to provide high-frequency performance, process variation tolerance, and low power consumption (Fig. 4.5.1). Performance and process tolerance are provided by a 2-clock design: early and late clocks (ECLK, LCLK) create a soft timing edge, allowing limited cycle stealing. Power is reduced in low-power SEFs by internally gated slave latch clocks. The majority of flops (78%) are low-power, using high-performance flops only on timing-critical paths.

In contrast to leveraged power-optimized CPU designs [2,4], Bulldozer's ground-up design requires co-development of power efficiency, timing, and functionality. Initially, micro-architectural power is optimized using a power-aware high-level performance model. Next, before schematic completion, the team tracks and analyzes RTL-based clock and flip-flop activity (a proxy for switching power) to meet clock gating goals. Finally, a new power model enables early mixed schematic/layout analysis of transistor-level power. This enables aggressive power optimizations while the implementation is still malleable. The result is a design with low power consumption for typical applications, making it well-suited to active power management and boost (Fig. 4.5.2).

The L1 caches are split, with I-cache residing in the instruction unit and a D-cache located in each load/store unit of the 2-cores. The 2-way, 64KB I-cache consists of an 8×2 array of 4KB bank macros, with 2 more arrays for pre-decode bits. Load/store area in the 2 cores is at a premium, so the D-cache uses a 4-way 16KB array with performance features described later in the paper. Both L1 caches use an 8T storage cell. The change from a 6T cell in 45nm to 8T in 32nm was required to improve low-voltage margin and read timing and to reduce power. Use of the 8T cell also eliminated a difficult D-cache read-modify-write timing path. Reads use a 2-level pre-charged local/super bitline structure with delayed-onset keeper, single-rail, full-swing signals, and glitch latches.

Several D-cache performance features reduce conflicts and power. First, micro-banking reduces read conflicts to the same rate as a previous 16-bank 64KB design by retaining 16 logical banks. Next, when both read ports access the same 128b word (a common occurrence), port-A data feeds both ports, saving power and avoiding conflicts (Fig. 4.5.3). Load/store unit power features include a static TLB with power filter to avoid TLB reads if the page has not changed, and extensive use of static wakeup CAMs to avoid retries.

Each core integer unit [3] contains a 40-entry out-of-order scheduler supporting single-cycle wake and pick of up to four instructions. To save wake array power, source/destination compares use a pre-decoded dynamic AND "match CAM" instead of a traditional mismatch CAM. The integer datapath supports 0-cycle result bypass to dependent instructions. To remove critical-path wire delay, the physical register file arrays and address generator (AGEN) incrementor are replicated. Most register file read bits are single-ended, but dual-rail reads for a few critical bits supply clocked data to the dynamic shifter.

The floating-point (FP) unit adds many features within the same basic pipeline as previous AMD x86-64 processors:

a) new multiply-accumulate functionality and instructions, including AVX, AES, SSSE3, SSE4.1, SSE4.2, XOP, and PCLMULQDQ;
b) dual-threaded support;
c) 4-wide instruction issue; and,
d) increased size of performance-critical structures [2, 4].

To minimize bus lengths, the datapath is split into high and low halves (Fig. 4.5.4), each acting on 64 bits for AVX/SSE/MMX or 80 bits for X87 operations. Four fully pipelined FMACs each support one EP/DP or two SP ops per cycle with 5-cycle latency to dependent ops. Optimized algorithms and dense, hand-placed datapaths minimize area, reduce cap and RC delay for timing, and reduce power. Extensive fine-grain clock gating allows FP to achieve an idle active power of less than 2% of peak active power.

A 160-entry 10R/6W FP register file (FPRF) reads 2 or 3 sources for 4 instructions, and writes up to 6 results, per cycle. The FPRF is split horizontally into two 10-bank arrays of 91 and 73 bits respectively. Each array aligns with a split FP datapath, and differs significantly from the replicated integer unit RF array design [3]. A flexible 2 and 3-level read bitline structure is used to reduce power and latency across the large data arrays (Fig. 4.5.5).

The cache unit (CU) contains L2 data TLB, and L2 cache interface datapath and controls. The CU and other units have large, complex random-logic control blocks that required the use of design automation techniques for efficiency. The complex logic and aggressive 12 to 14 gates useable between flops requires a variety of approaches to meet timing. The use of "bounds" statements guides structured logic placement without forcing explicit locations. Route estimation during cell placement models a metal stackup in which dense lower layers are up to 20× more resistive than higher, fast layers. Many arrays also rely on design automation or reuse. Across the core, many queues and buffers are built from compiled array structures and the larger CU memory arrays for the TLB and prefetch history tables are copies of I-cache and I-TLB arrays.

The 2MB L2 cache array has a 64B line size and 6-cycle internal pipeline. The 0.258μm^2 6T bitcell comprises 128KB slices of eight 16KB data array macros each. Figure 4.5.6 shows each data array macro uses a traditional sense amp with 128 rows per bitline and supports both row and column repair. The L2 data array macro contains four banks of sub-arrays with two-cycle throughput, allowing a full cycle for read or write and a full cycle for pre-charge. As shown in Fig. 4.5.6, wordlines are active during cycle 1. Early in cycle 2, the isolation PFETs turn on, passing differential from the main to the sense bitlines. After a programmable self-timed delay, the sense enable fires, turning off isolation PFETs and word lines. The sense enable is active for the rest of cycle 2, allowing the sense amp to latch the output data. A column circuit interlock prevents isolation PFETs from turning off before the sense enable fires. Interlock logic also disables isolation PFETs during the first cycle of the access while the sense amp bitlines are in pre-charge.

References:

[1] Butler, M. "Bulldozer: A new approach to multithreaded compute performance." Hot Chips 22, August 24, 2010.
[2] Jotwani, R. et al. "An x86-64 Core Implemented in 32nm SOI CMOS," *ISSCC Dig. Tech. Papers,* pp. 106-107, Feb. 2010.
[3] Golden, M. et al. "40-entry unified, out-of-order scheduler and integer execution unit for the AMD Bulldozer x86-64 two-core CPU module", *Accepted to ISSCC 2011.*
[4] Golden, M. et al., "A 2.6GHz Dual-Core 64bx86 Microprocessor with DDR2 Memory Support," *ISSCC Dig. Tech. Papers,* vol. pp. 325-332, Feb. 2006.

 978-1-61284-302-5/11/$26.00

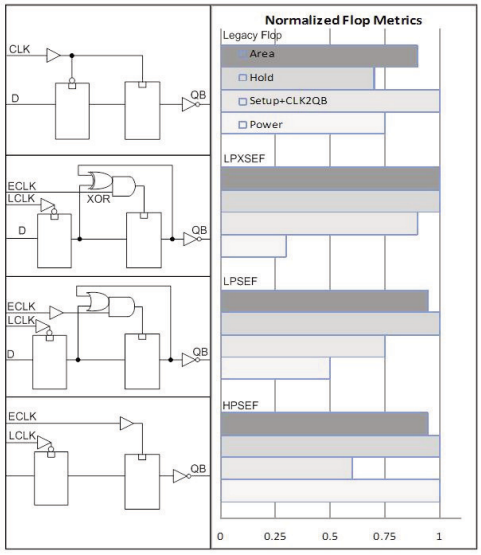

Figure 4.5.1: Normalized Soft-Edge Flop (SEF) metrics (lower value is better).

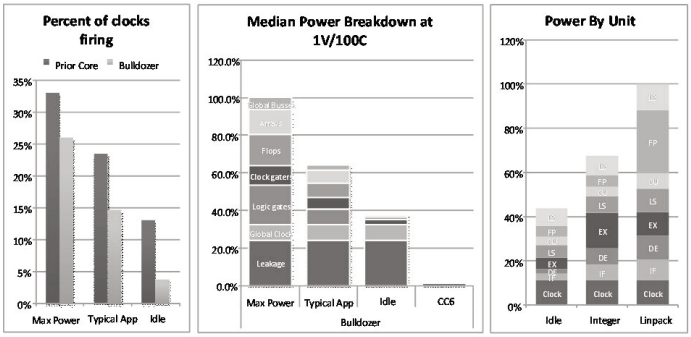

Figure 4.5.2: Bulldozer Power Consumption.

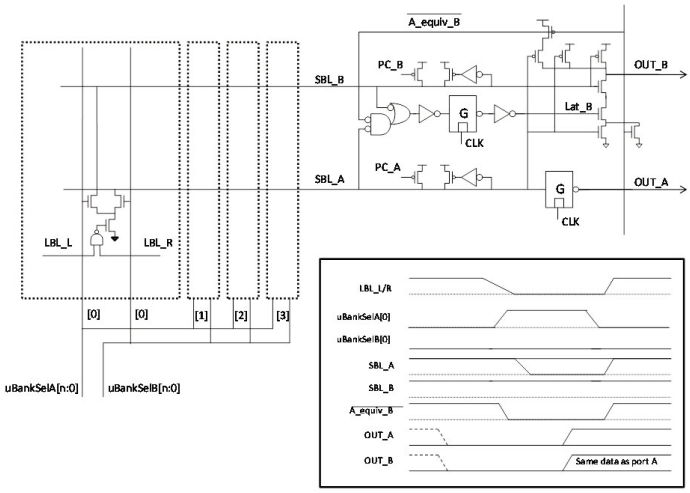

Figure 4.5.3: L1 D-cache Half-bank Output Latches.

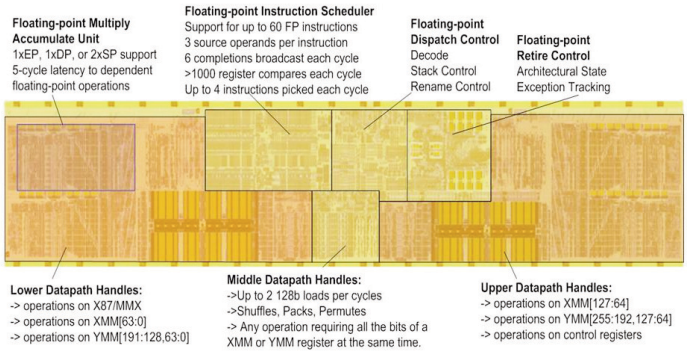

Figure 4.5.4: Floating-point Unit Floorplan.

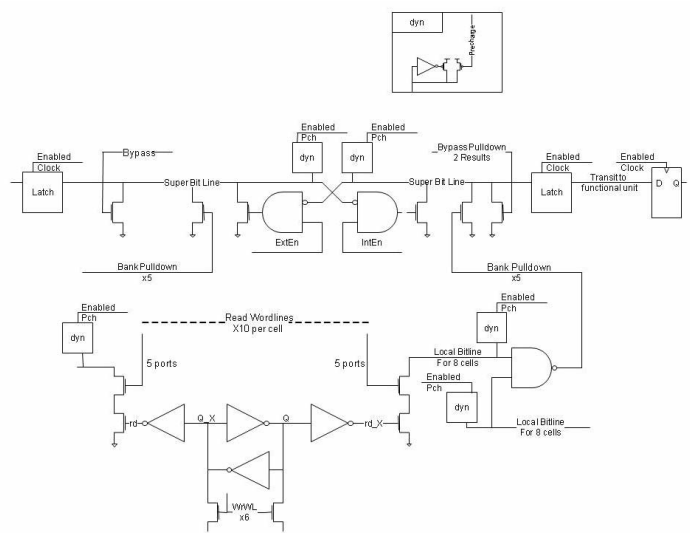

Figure 4.5.5: FP Register File Read Structure and Bypass.

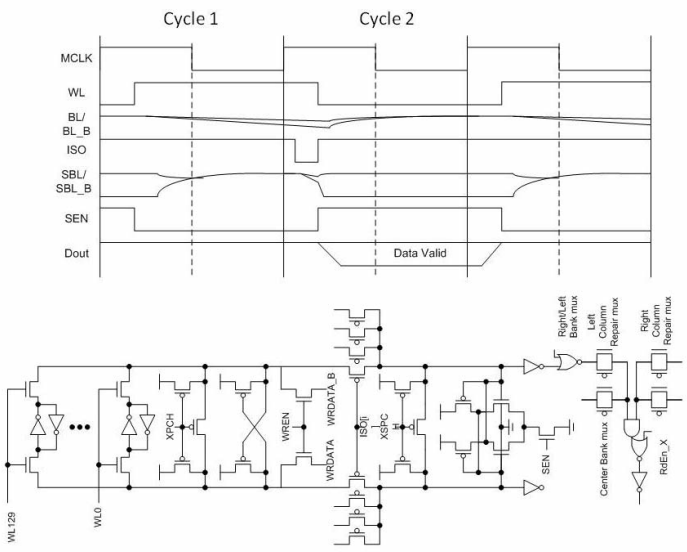

Figure 4.5.6: Cache Macro Read Datapath and Timing Diagram.

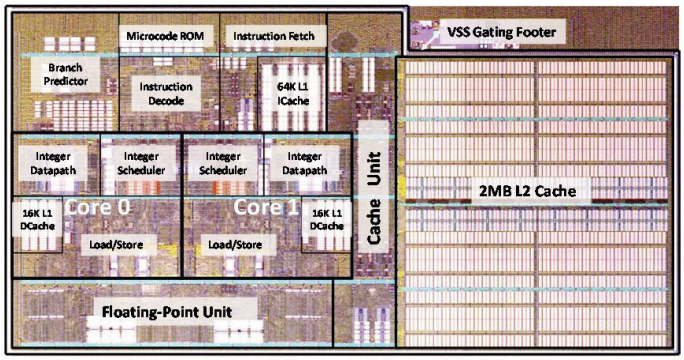

Figure 4.5.7: Bulldozer 2-core Processor Module Die Photo.

978-1-61284-302-5/11/$26.00